



The International
Fragrance Association

Member Associations
Members of IFRA Committees
cc: RIFM
Members of the JAG
Downstream User Associations

17 November 2025

Information Letter 1220

Subject: Summary Report – Proficiency Test on Methyl-N-methyl Anthranilate in Essential Oils

Related ILs: None

Dear Colleagues,

The IFRA Analytical Working Group (AWG) would like to inform IFRA members and stakeholders about recent work to demonstrate the effectiveness of analytical approaches used for quantifying Methyl-N-methyl anthranilate (CAS RN number 85-91-6) in an essential oil.

Introduction

A proficiency test on Methyl-N-methyl Anthranilate in essential oils was initiated by IFRA and commissioned to Fapas®, a proficiency testing provider which is part of Fera Science Ltd, UK. The final official report was published to IFRA and participating laboratories in June 2024.

Proficiency testing: outline

A proficiency test (or proficiency testing, PT) is an essential element of laboratory quality assurance. Test samples, replicating real-world analytical samples, are distributed to participating laboratories who are required to return their test results to the PT provider. In return, they receive an assessment of performance (i.e., overall laboratory competence) which compares their laboratory to all the others taking part. This provides those stakeholders with valuable insight into how easy or difficult an analysis has been and what implications that has for routine analysis services being offered by those laboratories to the industry. PT is at its most effective when undertaken as an independent exercise, commissioned to a professional PT provider.

Objective of PT for Methyl-N-methyl anthranilate

IFRA | THE INTERNATIONAL FRAGRANCE ASSOCIATION

IFRA Headquarters Cours de Rives 11, 1204 Geneva, Switzerland
IFRA Operations Avenue des Arts 6, 1210 Brussels, Belgium

+32 (0)2 214 2060 | ifrafragrance.org

The test analyte, Methyl-N-methyl anthranilate (CAS RN number 85-91-6) is restricted by IFRA ([IFRA Standard](#)) as well as for use in cosmetic products in Europe under Entry 323, and in the ASEAN Cosmetics Directive.

The Analytical Working Group (AWG) of IFRA recognizes that there are a variety of methods in existence for the analysis of this material in raw materials and producing a variant of these would serve no useful purpose. However, as a means of validating those methods already in use, and to establish a profile of the effectiveness of these methods, a PT provides a very useful insight whilst also not disturbing the participating laboratories work schedule unduly.

The test parameters were carefully considered and a spike level commensurate with an agreed 'typical' dose level in the test matrix. Three types of test materials were prepared from an essential oil, previously known to be free of the target analyte. The bulk essential oil was divided into three portions (labelled A, B and C) for spiking. The target analyte Methyl-N-methyl anthranilate (purity 99.17%, independently determined) was spiked into the essential oil samples labelled A (0.2% w/w) and C (0.7% w/w). Sample B essential oil was not spiked.

The start date was 20 February 2024 when aliquots of each sample type A, B and C were distributed to the 14 registered participants. The participants did not know in advance which sample type was spiked at what concentration, only that each sample had to be analysed for the target chemical Methyl-N-methyl anthranilate. In order to increase the number of data points and, therefore, the robustness of the data, each participant was required to submit two independently analysed results. The deadline for submission of results by participating laboratories was 9 April 2024.

Note: Re: performance assessments from a PT

A result is assessed in the form of a z-score which provides an indication both of the direction and magnitude of bias of that result:

- A positive z-score shows over-estimation of the result
- A negative z-score shows under-estimation
- A z-score of exactly 0.0 indicates no bias.

The further the z-score deviates from 0.0, the greater the magnitude of that bias. For the purposes of this IL publication, the 'satisfactory' range of z-scores is -2.0 to +2.0. A z-score lower than -2.0 or higher than +2.0 indicates either a warning signal or an action signal, depending on its magnitude. The statistical concept behind this is described more fully in the Fapas® Protocol Part 1 (reference 1).

Results and data analysis

The submitted results for the spiked samples showed a near-typical distribution for such analyses though, it was noted, this variation was above what is typically considered the 'norm' for the accepted Horowitz model (see Footnote below for a more detailed explanation).

From the 14 registered participants, 13 returned results but not necessarily for all the samples provided. In total, there were 23 sets of results returned out of a possible 28 (82%). Data for each sample type A, B or C were analysed separately.

- Sample A was a low level (0.2% w/w) spiked sample and the consensus of participants' results was 0.227% w/w, in good agreement with the theoretical value. Performance assessments were provided for each result, however, only 14 of 23 results (61%) were within the 'satisfactory' range of performance.

- Sample B was not spiked and participants returned results indicating either not detected or only trace levels of Methyl-N-methyl anthranilate. Therefore, no performance assessments were issued for this effectively blank sample.
- Sample C was spiked at 0.7% w/w and the consensus of participants' results was 0.814% w/w, in good agreement with the theoretical value. Performance assessments were provided for each result, however, only 16 of 23 results (70%) were within the 'satisfactory' range of performance.

The results show a good agreement, overall, with the spike value targets with a positive bias being seen in the data overall across both spike levels, this bias being between approximately 10 and 20% from the low to high spike value.

This could arise from individual variance in the methods used within each laboratory (i.e., from the nature of the GC column used, the GC method itself applied to the sample, single or dual phase analysis – polar/apolar chemistry for example, the use of FID and/or MS as the detection method and the data interpretation model used to generate the final reported result). All these may influence the final data reported but the PT also shows that, even with these potential variances, the final data is generally acceptable and fit for use by the industry.

In addition, whilst participation for the group was generally high with 13 of 14 participants returning results, these were not 100% complete and, therefore, this may have affected the overall interpretation and outcome of the PT.

Conclusions

This PT has demonstrated that the analysis of Methyl-N-methyl anthranilate in an essential oil is not straightforward. The distribution of results from multiple participating laboratories are non-Gaussian, more variable than the standard model predicts (see below) and that there are indications of a general positive bias of these results compared to the known preparation of the test materials.

While not ideal, this does not pose a significant issue for the industry as a positive bias means there is likely to be an element of safety in the data being reported. It is the participants' responsibility to investigate how this bias arises in their own analytical methods and to take the appropriate actions to correct these as outlined in the objectives of the PT model noted above. Further proficiency tests, combining greater participation of laboratories equipped to undertake the analysis together with a greater variety of target analytes and concentrations in the test materials, would confirm the effectiveness of these actions.

Footnote: Interpretation of assessments

What is 'satisfactory' or 'acceptable' performance in the context of this PT? It's highly unlikely that absolutely identical results will be submitted by all participating laboratories, i.e., there will be variation among results in the data sets. Therefore, what are the acceptable limits of this variation? What are the absolute values of results' magnitude that provide a z-score within the 'satisfactory' range?

In the field of analytical chemistry, there is a generally adopted model of acceptable variance, known as the Horwitz equation. This equation predicts the acceptable variance of analytical results, depending on the concentration of chemical being analysed (reference 1). The Horwitz model was first described in the early 1980s and it continues to be both used and verified as an appropriate model to this day.

In this PT, we observed that there was variation in the submitted results. However, the variation is far beyond what would normally be acceptable using the Horwitz model. In fact, we had to adopt two-times the variance predicted by the Horwitz equation in order to make

sense of the data. In other words, the disagreement between results in this PT is double that we would ordinarily expect in chemical analysis. Another way of looking at it is, the current 'satisfactory' range of ± 2 z-scores should really be in the range of ± 1 z-scores.

There was an additional feature noted with the data. Chemical analytical data should follow a normal distribution, i.e., a symmetrical bell-shaped curve centred on the assigned value (the accepted correct concentration of chemical). The observed data in this PT are not symmetrical or a uniform bell-shaped curve; they are not normally distributed data. The shapes of the distributions for spiked samples A and C are shown in Figures 1a and 2a (using statistical software to generate the kernel density plots, see reference 1). Even considering the non-normal distributions, the consensus of results for each sample A and C indicates a small (approximately 10%) positive bias.

To put some context into the shapes and spread of the results distributions, Figures 1b and 2b show what the ideal distributions would look like, for the same assigned values and axis ranges, assuming the data fitted the Horwitz model.

References

- (1) Fapas® Protocol Part 1 – Common Principles, www.fapas.com/technical-documentation

We hope that this letter will provide you with helpful information and invite you to contact Jennifer Dorts (jdorts@ifrafragrance.org) and Matthias Vey (mvey@ifrafragrance.org) in case of any questions.

Best regards,

IFRA

Figure 1a: Methyl-N-methyl anthranilate results distribution for sample A (observed)

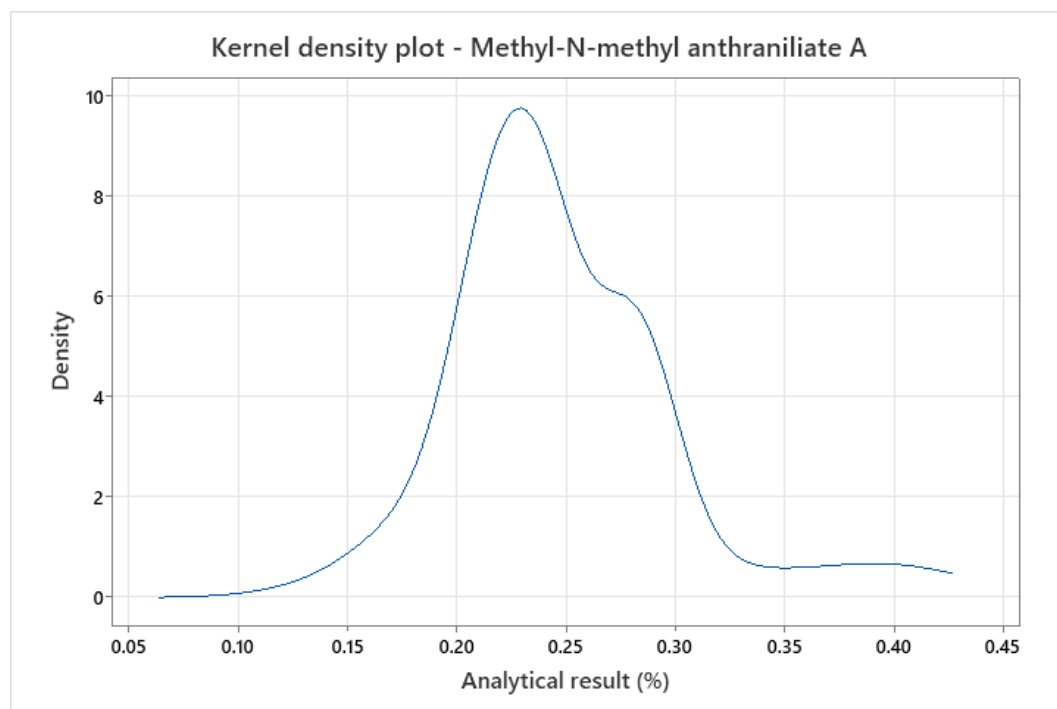


Figure 1b: Methyl-N-methyl anthranilate results distribution for sample A (ideal)

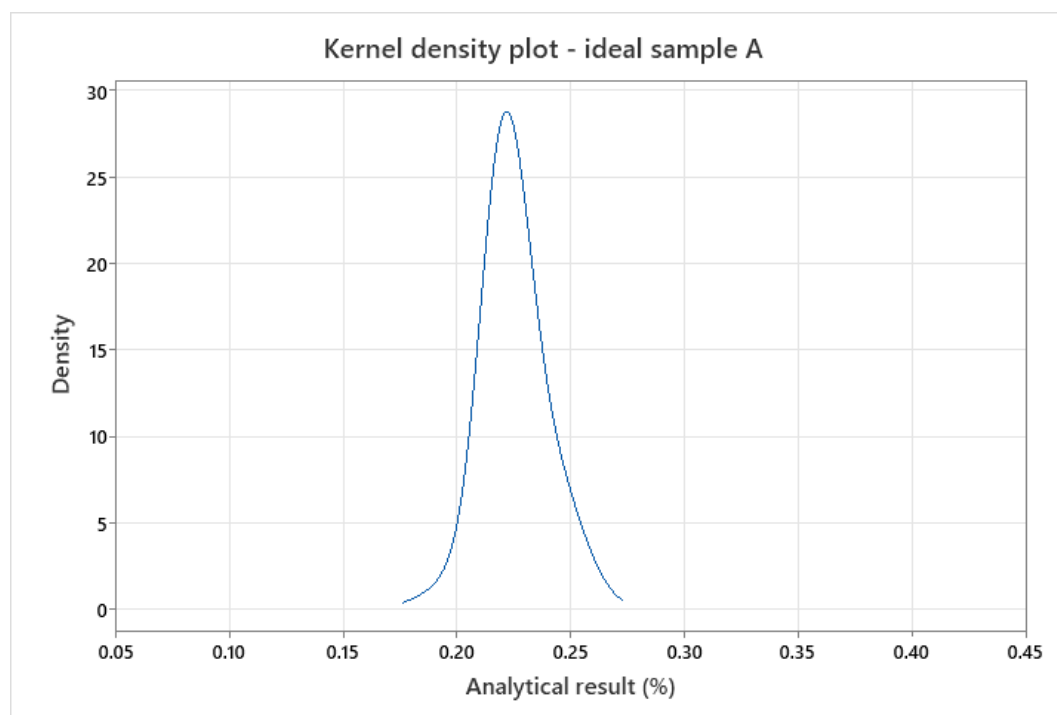


Figure 2a: Methyl-N-methyl anthranilate results distribution for sample C (observed)

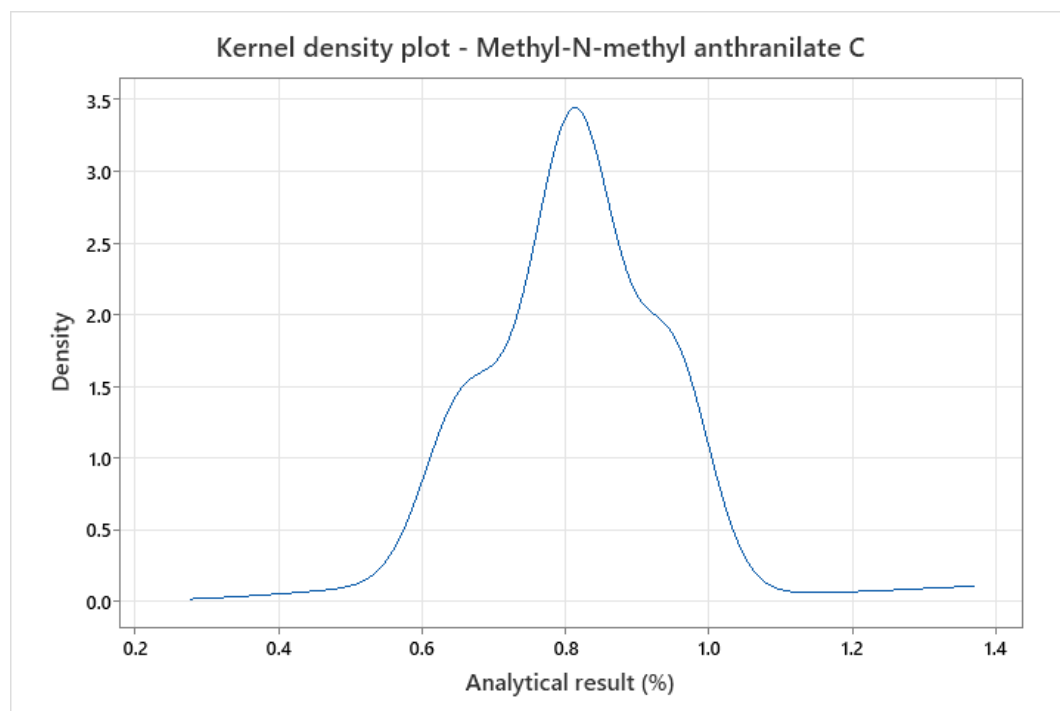


Figure 2b: Methyl-N-methyl anthranilate results distribution for sample C (ideal)

